Phenomenological Ontology and Supervenience

Andrea Zhok
Università di Milano
andrea.zhok@unimi.it

Abstract. The term 'supervenience' has acquired central status in contemporary philosophy in the wake of Davidson's discussion (1970) of mental e physical properties. We want to recover Davidson's original setting and to tackle the same problem by an alternative route, which resorts to a phenomenological approach. The notion of supervenience emerges at the crossroad between three theses: 1) the background of ontological monism; 2) the interpretation of monism as physicalism; 3) the dependence of mental events on physical events. 'Supervenience' bears the task to express simultaneously the ontological dependence of the mental on the physical and the apparent irreducibility of the former to the latter. In this framework, we are going, first, to grant ontological monism while rejecting physicalism. We will argue for an ontological framework that acknowledges some of the reasons of classical idealism, without abandoning a materialist framework. Secondly, we are going to show why intentional phenomena can never be regarded as epiphenomena. This move recommends a renewed methodological 'respect' for phenomena, consistent with the phenomenological tradition. Thirdly, by re-interpreting the old notion of formal causality, we will make room for an understanding of 'mental causation' that avoids both dualism and any violation of the principle of conservation of energy. In conclusion, we are going to argue that the reasons that originally motivated the introduction of 'supervenience' are inconsistent and that, in the theoretical context that we argue for, resorting to the notion of supervenience is either superfluous or misleading.

The term 'supervenience', in its contemporary acceptation, has become a philosophical term of art since its occurrence in Davidson's essay on *Mental Events*.[1] Here the reference to supervenience was introduced to capture a specific intuition: the idea that mental properties may depend on physical

1 Davidson 1970.

Metodo. International Studies in Phenomenology and Philosophy
Vol. 3, n. 2 (2015)
ISSN 2281-9177

properties while not being reducible to physical properties. By talking about mental properties that 'supervene' on physical properties it was possible to give expression to a form of dependence, which was akin to causal dependence, but avoided the *reduction* of mental properties to *physical* effects or causes.

From then onwards, the literature on 'supervenience' has considerably grown and has also become quite technical, in an attempt to find a formulation of supervenience that could formally capture the kind of non-reductive dependence that was originally at stake, and its logical variants. Yet, it is dubious whether such a formal direction of inquiry has positively contributed to clarify the original theoretical problem tackled by Davidson.

The following pages will try to recover the original Davidsonian question and to tackle it by a different route. More precisely, we will re-discuss the presuppositions shared by Davidson (and others) when resorting to supervenience, and we will do it by adopting a phenomenological approach. The reference to phenomenology is not meant to suggest a historical-philosophical confrontation. We will make use of some well-known features of Husserlian phenomenology, but we are not going to enter any subtle exegetic debate. We will argue that a phenomenological re-discussion of the conceptual constellation where 'supervenience' lies shows that the notion of supervenience is essentially *superfluous*, even if it does signal a crucial problem.

A final caveat about the organization of the argument: in order to effectively provide such a re-discussion we need to focus on a plurality of momentous issues, each of which would deserve a more thorough treatment than we are in a condition to provide in the present limits. Therefore, the argument that follows should be taken more like an edifice, counting on the support of a plurality of props, than like a chain, counting on the proven solidity of each of its links.

1. Davidson's 'Anomalous Monism' and Supervenience

Donald Davidson discusses the relation between the physical and the mental spheres by resorting to an apparent paradox, derived from the juxtaposition of three allegedly self-evident principles. The first one is called by Davidson the *principle of causal interaction* and says that «mental events interact causally with physical events».[2] The second one is called the *principle of the nomological*

2 DAVIDSON 1970, 208.

character of causality and states that «events related as cause and effect fall under strict deterministic laws».[3] The third one is the *principle of the anomalism of the mental*, which states that «there are no strict deterministic laws on the basis of which mental events can be predicted and explained».[4] At first glance the three theses seem mutually incompatible, since from the first and the second we should draw the conclusion that mental events also fall under strict laws, which contradicts the third thesis. However, Davidson rejects such incompatibility by arguing as follows:

> Causality and identity are relations between individual events no matter how described. But laws are linguistic; and so events can instantiate laws, and hence be explained or predicted in the light of laws, only as those events are described in one or another way. The principle of causal interaction deals with events in extension and is therefore blind to the mental-physical dichotomy. The principle of the anomalism of the mental concerns events described as mental, for events are mental only as described.[5]

If should we focus on a critical analysis of Davidson's argument, many theoretical details should be added concerning his characterization of notions like 'law' and 'language'. But for our purposes we do not need to be concerned with anything more than a general outline of his thesis, which can be rendered as follows. In a *monistic* framework, mental and physical events interact. *Events* as such belong to the extensional sphere. Events are the ontological substance to which we can give descriptions in physical or mental terms. At the ontological level, that is at the level of events *per se*, *causal interactions* take place. On the contrary, the subsumption of events under physical laws and the bestowal of meaning to them belong to the sphere of descriptions. Therefore, according to Davidson, at the ontological level we can have event causality between events which are described as mental or physical, but this does not imply that mental accounts should or could be subsumed under strict physical laws.

The most debatable aspect of this vision emerges when Davidson makes explicit the kind of monism that he is committed to. In agreement with mainstream contemporary consensus Davidson assumes that the only plausible kind of monism is a physicalistic monism. This implies, as he says, that at the ontological level *all events are physical*. But this produces considerable theoretical tension, since the attribute 'physical' enters at two levels of his

3 DAVIDSON 1970, 208.
4 DAVIDSON 1970, 208.
5 DAVIDSON 1970, 215.

account: at the *epistemic* level of descriptions, and at the all-comprehensive level of *ontology*. On the one hand Davidson is treating physical attributes as features dependent on human judgment, on the other, however, nature in itself (regardless of anybody's judgment) is said to be 'physical'. This duality is hardly transparent. How can Davidson say that his ontology of events is expressly 'physical' and what is meant by that, if not a judgment that depends on a bestowal of meaning?

It is precisely in this troublesome passage that the question of supervenience comes to our attention:

> Although the position I describe denies there are psychophysical laws, it is consistent with the view that mental characteristics are in some sense dependent, or supervenient, on physical characteristics. Such supervenience might be taken to mean that there cannot be two events alike in all physical respects but differing in some mental respect, or that an object cannot alter in some mental respect without altering in some physical respect. Dependence or supervenience of this kind does not entail reducibility through law or definition.[6]

Supervenience is here introduced in its intuitive formulation, so that there cannot be a difference in the mental (supervenient) sphere without a difference in the physical (subvenient) sphere. Although supervenience is no form of causality, for Davidson this kind of *dependence* is actually rooted in a 'causal' ground. Ultimately the supervenient sphere (whatever it may be) depends on the ontological level where 'event causality' takes place. This is the extensional ground where all 'causal interactions' take place even when no 'causal law' is available.

In the Davidsonian account we find two acceptations of 'causation': the first one concerns bare events (expressed by the principle of causal interaction), while the second one concerns the causal laws under which the instantiations of event causality can be subsumed (captured by the nomological character of causality). The second one is supposed to support the possibility of *reduction* of the mental to the physical, whereas the former one should be neutral on this point. Thus, in order to clarify sense and possibilities of supervenience we shall try to deal separately with 'causation': 1) from an *ontological* perspective and 2) under the prospect of *reduction*.

6 DAVIDSON 1970, 214.

2. Supervenience, Causation, and Monism

In the Davidsonian framework, and in most of the relevant debates on supervenience, supervenient properties are ultimately taken to supervene on physical matter as ontological basis. Most contemporary discussions assume a monistic background as the only ontological stance worth considering, and there 'monism' is usually a synonym of 'physicalistic monism' (or 'materialism').

2.1 The Reasons of Idealism and their Limits

The best argument that can be provided in support of monism, as far as we can see, can be derived from the aporias that its main alternative, Cartesian dualism, gives rise to. The defining character of substance is autonomy (*quod in se est et per se concipitur*). If there were more than a single substance, we would have only two possibilities: either different substances do not 'casually' interact or they do. If they do not, this would make of them something like parallel universes: no event pertaining to substance 1 could have effects of any kind on events concerning substance 2. There could be neither actions nor perceptions to bridge between substances and consequently, whatever the 'substance' of our pertinence, we would never be in a condition to know anything about the other one and our condition would be epistemically and practically equivalent to monism.

If on the contrary, we grant that two substances 'causally' interact, then it becomes unclear why we should conceive of them as *independent* substances, since they must necessarily have 'something in common', in order to interact. Thus, if we grant, as Davidson does, that mental and physical events causally interact, then we have good reasons to conclude that they do it on a monistic background.

That said, it is appropriate to remind that in historical terms monism took at least two main embodiments: idealism and materialism. No doubt, idealism is nowadays rather out of fashion. In part idealism has been marginalized for reasons that have little to do with decisive arguments. For instance, since classical idealism was rather hostile to natural sciences, the historical success of the latter spelled the decline of the former. Now, let us maintain for a moment that idealism is a plausible form of monism. We can immediately see that its application to a notion like supervenience generates some remarkable conceptual metamorphosis. If we take supervenience to be defined, in its most intuitive formulation, by the slogan "no supervenient difference without a subvenient difference", in the context of idealistic monism we could produce

the following true sentence: "there is no physical difference without a mental difference". This claim is empirically no less plausible than the (more widespread) inverse one. Indeed, whenever we grasp a physical difference this naturally involves the occurrence of a mental difference in the 'grasping' (thinking or perceiving) subject.

In a materialistic framework when we mention a 'physical difference' we assume that physical differences subsist per se, regardless of any subjective contribution, and therefore regardless of mental events: under these premises the mentioned case of 'idealistic supervenience' remains invisible. But, are there good philosophical reasons for granting that 'physical differences' subsist regardless of subjectivity? In what sense?

One of the basic arguments of idealism is that for anything that we can mention or argue for, for any scientific fact and any observational report, at any level of detail, there is always a subjective (mental) occurrence, and therefore the assumption that facts, events, things, etc. subsist in the absence of subjectivity would be sheer dogmatism. Yet, there is a questionable conclusion that idealistic monism used to draw from that observation, that is, the conclusion that Being (reality) must *depend for its existence* on subjective activity. This is the core of so-called *subjective idealism* (Fichte, Berkeley), which is also the best-known (and the least tenable) of the historical forms of idealism. But what is wrong with subjective idealism? If we put aside the appeal to cheap commonsense and mainstream consensus, this point can be settled, as far as we can see, only from the following perspective.

It is true that for any perceived quality, any meaningful thought, any attribute of things, etc. something must happen in the sphere of *intentional*[7] acts (mental events): there is no perceived redness without the specific reactivity of eyesight, there is no recognition of truth without a judgment, etc. And in this sense the set of mental events is at least as vast a set as the set of physical attributes (actually, it could be easily shown that it must be *more comprehensive*). Nevertheless, this does *not* imply the ontological *priority* of intentional acts on Being. That is, there is no good reason to claim the dependence of Being on intentional acts. This point can be motivated on strictly phenomenological grounds.

The first thing to notice is that there is a sphere of experiences, pre-eminently sensuous experiences, where intentional contents manifest themselves as something that does *not* depend on subjective activity. When we perceive, we

7 From now on, and throughout the text, we will use the terms '*intentional*' and '*intentionality*' exclusively in the phenomenological sense of mental power to *mean*, *represent*, *stand for* and *be about* objects, regardless of whether such objects are spatio-temporally existent or not.

can be *surprised* by what occurs and our apprehension of contents appear to follow cues that are *received* and *discovered*, not produced. This is what in Husserlian language has been labeled *transcendence* (*Transzendenz*). *Transcendence* in this sense is the sphere of what appears to consciousness as given independently from conscious activity, that is, as being beyond (trans-) conscious activity.[8] Consciousness (mind) manifests itself not just as activity, but also as reactivity and passivity.[9]

One could object that this notion of 'transcendence' refers just to how things appear, not to how they truly are. However, in a phenomenological framework, *experience* in its most comprehensive sense is *appearance*, and experience is the basis of all truthful knowledge. Phenomenological description (under *Epoché*) is a description that prima facie refrains from tackling the questions of reality and causality. No claim is made about the status (real or delusional, original or secondary) of the relevant appearances. However, any theory about mind and matter, cause and effect, reality and illusion must be rooted in an experience that, necessarily, does not know yet what mind, matter, reality, illusion, etc. are. Thus, phenomenological description under *Epoché* gathers and gleans from the most original source of evidences, on whose basis we can eventually argue for a specific understanding of reality, causality, truth, etc. Therefore, in order to defy evidences drawn from a phenomenology of *perception*, we cannot just appeal to theories that tell us what reality versus 'mere apperance' is: we must appeal to evidences at the same experiential level.[10]

Here, the notion of transcendence captures the phenomenal ground that supports our preference for a materialistic monism. Transcendence signals that our primal experience involves a sphere of intentional acts, which manifest themselves as not 'sovereign', but as 'cued' by a 'transcendent otherness'. This is the phenomenon that motivated some strained theoretical contrivances of subjective idealism, like Fichte's "unconscious activity of the absolute ego" and Berkeley's reference to a ubiquitous God's activity. They appear as *ad hoc* arguments to explain away the evidence, disturbing for subjective idealism, of irreducible transcendence.

2.2 Spinozian Ontology or Physicalism?

However, if we reject subjective idealism, this does not automatically strand us on the shores of physicalistic materialism. By recognizing the subsistence of a

8 Hua II, 72; HuaMat VIII, 52.

9 HuaMat VIII, 42.

10 Phenomena are to be originally apprehended "as what it is presented in being, but also only within the limits in which it is presented there" Hua III/1, p. 51.

transcendent sphere we may be inclined to embrace some form of materialistic monism, since we are justified in saying that conscious activity is not sovereign and that there is a transcendent 'matter' that can 'impose itself on us'. Still, nothing is said yet about the nature of 'material properties'. We can say that there is an independent 'efficaciousness'[11] of events in the world, but we have little to say about the essence of ontological events. This picture is consistent with what Davidson argues: event causality is supposed to take place regardless of any description of it in mental or physical terms. From this point of view the Davidsonian account seems to have a Spinozian flavour: the ontological substance anonymously subsists, while mental and physical attributes describe the substance without ever exhausting its nature. Yet, such a vision would be no more materialistic than idealistic, and here is where Spinoza and Davidson part ways. According to Davidson the ontological substance must also be regarded as physical: all events are physical. And what about event causality? Does it deserve to be called 'physical'? On this point Davidson is less clear than one might wish, but he does share the essential tenet of physicalism, that is, the idea of the causal closure of the physical world (if a physical event has a cause, it must be a physical cause). Indeed, his idea that all events are physical, while not all events are mental, depends precisely on this assumption: there are many events that affect the mental sphere while not being mental, and this means that the sphere of mental events is not closed,[12] whereas the sphere of physical events is supposed to be causally closed.

This is a crucial passage. Actually, saying that the sphere of mental events is *not* closed is improper, or at least unclear. If it just means that 'events happen' without being initiated or governed by conscious activity, this is undeniable. We can also grant that there are good reasons to believe that trees are falling in the wood without anybody perceiving it. But if this approach wants to support the pre-philosophical idea that there is a big world out there, where consciousness (mind) is bound to occupy a tiny section, this is a hasty, and ultimately untenable, conclusion. It should be superfluous to remind that this very picture of the big world, within which mind would lie as a minor subset, is itself a *mental event*. It is easy to see that conscious *activity* is limited, and it is also easy to grant that 'conscious behaviours' and 'conscious beings' are circumscribed sets among worldly items, but this has nothing to do with

11 We will use across the article the expression 'efficaciousness' to express a minimal notion of 'causality', which is not committed yet to any specific interpretation concerning *how* efficacious relations in nature must unfold (*e.g.*, over time or instantaneously; by contiguity or at a distance; with perfect uniformity or not, etc.).

12 DAVIDSON 1970, 224.

detecting the 'borders of consciousness'. Thus, one thing is to say that something transcends *conscious activity*, quite another to say that such 'transcendent something' is the 'physical universe'. 'Physical', as much as 'mental' are *predicates* in the sense of the Kantian remark that 'existence (*i.e.*, 'transcendent otherness') is not a predicate'.

2.3 Kim's Argument of Causal Exclusion

Now, to claim that reality (*i.e.*, Being, ontological substance) is 'physical' is neither obvious nor easy to argue. What does here 'physical' mean? When physicalists use the term 'physical' they do not merely refer to 'nature' (*physis*), but to nature as it is accounted for by physical science (or at most by the 'natural sciences' and pre-eminently by physics).

We can outline the background of physicalism and its meaning by completing Davidson's argument with Jaegwon Kim's critical remarks. Kim recalls that the argument for anomalous monism starts from the assumption that mental events can cause physical events. Thus, let us assume that mental event M causes physical event P. In the physicalistic framework that Kim shares with Davidson, we must grant that physical event P must have a *physical* cause. This derives from the "*principle of causal closure*", according to which «[i]f a physical event has a cause at *t*, then it has a physical cause at *t*».[13] However, if we grant that event P is caused by physical event P*, and if we remember that, by hypothesis, P is caused by mental event M, it seems that we have to do with an instance of overdetermination, where there are two sufficient causes for the same event. But causal overdetermination is hard to grant except as an occasional accident. At this point Kim reminds us that in the Davidsonian picture all events are physical and therefore also all mental events are in a sense physical. This is precisely the point where supervenience enters the stage: mental features 'supervene' on physical features. But at this point, it seems natural to draw the conclusion that we do not have separately to do with a mental and a physical sufficient cause, but that they are one and the same. As all events are, at their core, physical, all instantiations of event causality must be physical as well. According to Kim this is enough to show that mental events and mental properties may well subsist as descriptions, but *without any causal power of their own*: they are epiphenomenal. And this dispatches Davidson's argument for anomalous monism by denying the first of its premises: there is no causality from the mental sphere to the physical one.

Kim's argument is apparently robust and it certainly manages to expose

13 KIM 2005, 15.

Davidson's ambiguous use of 'physical'. But when it comes to clarify what is meant by 'physical', Kim does not seem to fare better. He thinks that the 'physical' prima facie includes all entities and properties of *basic physics*. Secondarily this set must be completed with any entity aggregated out of or composed by physical entities, and any second-order property defined over physical entities.[14] But this 'definition' provides just a contingent set of entities, without explaining what its essential boundaries amount to. This definition of the physical remains tied to what is contingently part of the historical reality of 'basic physics'. This does not seem to be a progress beyond the appeal to a contingent consensus.

2.4 Papineau's Definition of the Physical Realm

David Papineau subscribes to Kim's argument and tries to provide a more satisfactory account of what should count as 'physical'. The gist of the reference to the physical sphere is the idea of the *closure of the physical realm*.[15] Papineau observes that current physics cannot be plausibly regarded as complete and that no list of all fundamental forces and particles responsible for physical effects would provide anything like a closed all-embracing field,[16] to be opposed to the allegedly open field of mental events. Actually, the problem of the closure of the physical realm splits into two questions: one concerning the definition of what is 'physical' and the second ensuring that what is physical is always cause and effect of something 'physical'. Papineau tackles these points as follows. He first tries to determine what is 'physical' in *negative* terms, that is, not by spelling what should essentially belong to anything 'physical', but by defining what should be excluded by it. And secondly, he tries to define the causal closure of the physical realm with reference to the principle of the *conservation of energy*.

The first argument assumes that we could define 'physical' anything that is *not* identifiable through mental characteristics, thus, «we may not know enough about physics to know exactly what a complete 'physics' might include», but whatever it includes «it will have no ineliminable need for any distinctively mental categorizations».[17]

Yet, this is a baffling excogitation, since it should be obvious that more or less all of physics is composed of concepts, inferences and equations, which are precisely "mental categorizations". Probably Papineau would like to refer just

14 Kim 1997, 294.
15 Papineau 2002, 9.
16 Papineau 2002, 41.
17 Papineau 2002, 41.

to all entities a parte objecti, regardless of how they are objectified, and then to exclude from the set all the 'bearers of psychological features'. But this is in turn a *semantic* distinction between fields that do have an "ineliminable need for mental categorizations". This objection has nothing to share with the idealistic claim that there is *nothing beyond* mental categorizations. Nature may well have transcendent existence but we are nevertheless not in a position to establish that the features and properties that we attribute to nature (starting with its subdivision into mental and physical 'regions') are *independent* of the mind. Natural distinctions may be empirically justified, but criteria and experiences that provide such justification are ineludibly dependent on intentional acts (even if the extent and form of such dependence may remain unsettled).

If we come to the second argument by Papineau, we see that the causal closure of the physical realm amounts to a vindication of the principle of the conservation of energy.[18] There is no doubt that this is an authoritative principle in physics and we have no interest in challenging it, but the point is whether this principle establishes something like a specific 'essence of physical causation' or not. If the principle of the conservation of energy defined a 'way of producing effects' peculiar to physics, then we would have an interesting argument in favor of a specifically *physicalistic* ontology. If the conservation of energy implied that natural causes must have 'such and such' features, we might be forced to drop the claim, for instance, that also desires can truly cause the motion of limbs. Under these circumstances we should then say that mental appearances supervene on physical events, whereas true efficaciousness belongs only to the latter.

But the principle of the conservation of energy, regardless of its possible limits, does not say anything about *how* energy can be possibly deployed. In modern physics there is an updatable list of the forms that energy can take (gravitational, electromagnetic, nuclear, etc.), but essentially energy is just 'what can perform a work', and if we find new ways in which work can be performed, then room is made for new features of energy. Actually, although there are authors that interpret causality in the light of the idea of a conserved physical quantity[19] there is no orthodox theory of causality in physics. There is, in fact, only one point where the acceptance of the law of conservation of energy could conflict with the idea of 'mental causation': this would happen, if the spontaneity of mental causation implied the *production of energy ex-nihilo*. But, as we are going to see, this is an unnecessary claim.

18 Papineau 2002, 247-249.

19 Fair 1979; Salmon 1998; Dowe 2000.

2.5. First Intermediate Conclusion

The notion of supervenience came forth at the conceptual crossroad between three theses:

1) A background of *ontological monism*, such that mental events (or properties) and material events (or properties) causally interact and belong to the same 'substance'.

2) The interpretation of monism as *physicalism* (or mainstream *naturalism*), which implies that the one substance of reality is 'physical' (or anyway is properly qualified by the categories of the science of nature) and that all 'ontological efficaciousness' develops according to the model of physical laws.

3) The idea of an essential *irreducibility* of mental properties and events to physical descriptions.

Supervenience was in charge to express at the same time the ontological dependence of the mental from the physical, whose ontological priority was insured by thesis (2), and the apparent irreducibility of the mental to the physical, claimed by thesis (3). In Davidson's terms the supervenience of the mental was meant to grant mental causation. But Kim's argument of causal exclusion is apparently inescapable: if the mental produces real effects, this can be the case only because of its physical embodiment; and symmetrically, if there were mental aspects irreducible to their physical embodiment, they could be nothing but epiphenomena.

However, we have seen that the reasons that support monism are wholly separate from the ones that lead to physicalism. We have reasons to concede the transcendent subsistence of a unitary ontological background where all efficacious relations take place. But we have no reasons, apart from contingent cultural congeniality, to accept that this ontological background is 'physical'. This means that we might still want to say that mental properties *supervene*, yet we are not entitled to say that they supervene on *physical* properties: at most they supervene on transcendent featureless Being. Of such Being we are entitled to say just that it has 'efficaciousness' of some kind: it produces 'effects'. This implies that we are not bound to acknowledge any specific form of causality (like the one that, according to Davidson, is captured by strict physical laws) and therefore we are not forced to the dilemma either to reduce 'mental efficaciousness' to 'physical' (efficient) causality or to accept mental epiphenomenalism.

Still, the notion of the supervenience of the mental sphere on the physical one is also an attempt to deal with an intuition, the empirical justification of which is vast and growing. Nobody doubts that drinking a coffee wakes you up, taking a sleeping pill makes you somnolent, being hammered on a toe makes

you hurt, etc. And the more our neurophysiological knowledge grows, the more we learn how such and such material causes can induce specific mental states. Since in all these cases we can describe the relevant causes in physical terms, the question concerning how the elicited mental events are to be interpreted at an ontological level is compelling.

3. Supervenience, Reduction and Epiphenomenalism

As we saw, supervenience emerged in contemporary debate as a philosophical term of art in order to make room for a form of *non-reductive* dependence of the mental on the physical. The issue of reductionism is one of the most debated in 20th Century philosophical landscape and we cannot try here any comprehensive discussion of this theme. Thus, we shall keep our discussion close to what is strictly relevant to the present issue.

3.1. Can Kim's 'Functional Reduction' Yield a 'Conceptual Reduction'?

Reductionism is a thesis that can be discussed from a methodological, ontological or epistemic perspective. Classical reductionism[20] was an attempt to give epistemic plausibility to the assumption of a physicalistic ontology. Among the various criticisms that have been raised against Nagelian reduction, we want to mention just the one that Jaegwon Kim, himself a reductionist, reports. According to Kim, the classical model of reduction suffers from the fact that bridge laws, *i.e.* the laws that should enable to reduce the vocabulary of a theory T1 to the vocabulary of a second more comprehensive theory T2, are merely inductive (contingent) connections.[21] In this sense Nagelian reductionism did not yield any '*conceptual reduction*', where the truthful conceptual content of the reduced theory is comprehended in the reducing theory.[22] The point that this remark makes explicit is that the core aspiration of reductionism was to show that in the end all true theories should be conceptually *homogeneous*. All true theories should be translatable into the dominant language of the most basic and foundational theory, which was supposed to be 'physics'. But Nagelian reduction only manages to produce nomological ties between terms, that is: there is no translation of the *meanings* of the theory to be reduced (T1) to the reducing one (T2), but an equivalence is

20 NAGEL 1961.
21 KIM 2006, 194-195; 1992, 125-126.
22 Cf. KIM 1999, 7-8.

established only in the form of empirical correlations between events referred to by terms in T1 and T2 respectively. So, if pain and the excitation of C-fibers co-occur, we might be entitled to reduce pain to the excitation of C-fibers, even if many meaningful features of pain are missing in the concomitant neurophysiological description.

Kim believes, however, that this problem of classical reduction can be overcome. In this regard, he illustrates an alternative model of reduction called 'functional reduction'.

If we take *B* to be a domain of properties (phenomena, fact, events, etc.) that serves as the reduction base, while *E* is a property to be reduced:

> The reduction of property *E* to *B* involves three steps:
>
> Step 1: *E* must be *functionalized* – that is, *E* must be construed, or reconstrued, as a property defined by its causal/nomic relations to other properties, specifically properties in the reduction base *B.* (…)
>
> Step 2: Find realizers of *E* in *B*. If the reduction, or reductive explanation, of a particular instance of *E* in a given system is wanted, find the particular realizing property *P* in virtue of which *E* is instantiated on this occasion in this system;
>
> Step 3: Find a theory (at the level of *B*) that explains how realizers of *E* perform the causal task that is constitutive of *E* (i.e., the causal role specified in Step 1). Such a theory may also explain other significant causal/nomic relations in which *E* plays a role.[23]

We must recall that classical reductionism entered in crisis, among other things, under the functionalist challenge that the same property can be implemented by a multiplicity of different substrates. Kim's account of functional reduction grants that properties can be multiply realised, and requires, in order for reduction to take place, just to redefine in *causal* terms (operant at the level of base properties) the property to be reduced. The level of base properties is for Kim the physical level. Once the property has been thus functionalised, we should search for appropriate realisers at the physical level. Finally, the connection between the causal functionalisation and the discovered realisers should be supported by an appropriate theory that "*explains how*" the realisers perform the causal task defined by the proposed functionalisation.

23 KIM 1999, 10-11.

One of Kim's favourite examples of functional reduction is the reduction of genes to DNA: firstly, "genes" are defined in terms of their causal power to reproduce phenotypic traits across generations; secondly, the causal efficacy of DNA in carrying those traits from one generation to the following is discovered in the realm of physical properties; and finally, «we have a theory that explains just how the DNA molecule is able to perform this causal work».[24] This procedure is supposed to reduce the apparent heterogeneity between second-order properties (like "gene") and first-order properties (the physical instantiations of deoxyribonucleic acid). This kind of heterogeneity had not been really reconciled by classical reductionism, because of the merely inductive character of bridge-laws, whereas functional reduction is supposed to provide a *theoretical* bridge that *explains how* the macroscopic property of reproducing phenotypic traits in reality is (identical with) a physicochemical property of a nucleic acid.

If Kim's functional reduction turned out to be tenable, we could get a reduction of all general properties (functions)[25], among which mental properties, to physical properties, and more specifically we could have a reduction of the efficaciousness of the mental properties to the kind of causality that we are supposed to meet at the physical level. The latter conclusion amounts to the result that whenever a mental property seems irreducible to physical properties it is bound to be an epiphenomenon.

3.2 Can Intentionality be Reduced to Extensionality?

Now, let us tackle each of Kim's claims.

First, is it truly plausible that Kim's account manages to reduce functions[26] ('essences', 'meanings') to particular physical concretions? Functional reduction states that if we recognise that property P is essentially characterised by causal power C and if we discover a physical substrate S to which C can be attributed, then we can posit P = S.

24 KIM 1999, 10.

25 I regard in this context general properties and functions as synonyms. I take properties to be defined by their ability to bring about effects/outcomes endowed with the same identity; and I take functions to be defined by the capacity to issue into particular realizations endowed with the same identity.

26 I take the notion of 'function' to be intelligible only as an *intentional* relation. Functions must be able to bring about effects/outcomes that preserve a specific *identity*. In principle we can conceive of the effects of a cause without any identity to be preserved: they can go freely astray. On the contrary the realizations of a function are bound by definition to preserve a settled identity. The tie that connects premises and outcomes of a function must refer to the preservation of a common *meaning* or *essence*, and this is not intelligible without reference to intentionality. (See below, the account of efficient causality).

Yet, let us assume that P is "being somniferous", which translates into: P is endowed with the causal power C of producing sleep. Then, if we find a physical substrate S capable of causing sleep (*e.g.*, Phenobarbital), we should be entitled to say that P = S, that is, being somniferous equates with Phenobarbital. But this is clearly incorrect, since there are many other physical substrates endowed with the same causal power (*e.g.*, Diazepam).[27]

It has been sometimes claimed that this difficulty can be overcome by equating "being somniferous" with the *disjunction* of *all* sleep-inducing substances (either Phenobarbital or Diazepam or Lorazepam, etc.). But this equivalence is inacceptable because the *property* "being somniferous" refers to *infinite possible* embodiments of the sleep-inducing power, and never to *finite actual* embodiments. In fact, no property (or function, or meaning) can be considered equivalent to a list, not even an exhaustive list, of particular embodiments, because its *intentional* nature always exceeds any finite set of concretions (realisations, instances): it is always possible and legitimate to add a further entity covered by the *same* intentionality (*i.e.*, realising the same function).

3.3 What is Inherited Through "Causal Inheritance"?

It seems that the only way in which the ontological gap between the intentional sphere of properties (functions, meanings) and the extensional sphere of particular embodiments can be reduced is by denying true ontological status to the former. This is, indeed, what Kim proposes, by suggesting the ontological vacuity of functional properties. This point is argued by reference to the so-called "*causal inheritance principle*", according to which «[i]f a functional property *E* is instantiated on a given occasion in virtue of one of its realizers, *Q*, being instantiated, then the causal powers of this instance of *E* are identical with the causal powers of this instance of *Q*».[28] The causal inheritance principle essentially states that all ontological efficaciousness inheres in particular *physical realisers* and that such efficaciousness is simply 'inherited' by higher-order properties. In this sense, functional properties and in general second-order properties (among which mental properties) are considered void of any causal role different from the causal powers of their realisers (first-order properties, *i.e.* physical properties). It goes without saying that under these premises universals, meanings, functions (and mental properties among them) are not ontologically efficacious, or rather, they have no efficaciousness distinct

27 Cf. Block, 2007, 2.

28 Kim 1999, 16.

from the efficaciousness of their particular physical embodiments: as Kim states, the only justifiable role left for functional properties is epistemic, since they «may play a practically indispensable role in our discourse»,[29] but they have nothing to contribute at the causal level.[30]

It is clear that the reference to 'causal inheritance' is meant to identify a form of 'true causality' to be separated from what is only apparent causality, otherwise there would be no point in stressing the '*identity of causal powers*' throughout the causal process. Thus, we should say, for instance, that physical causes 'remain exactly the same' when they operate within a biological organism and when they operate outside of an organism. Yet, the question is: how are we going to decide whether this is truly the case? In order to do so, we should be in a position to establish what is the true essence of causality at the level of physical realisers, that is, the essence of 'physical causality'. But, as we have seen, we cannot rely on any obvious orthodox notion of physical causation, and, more important, we do not have any ground to limit ontological efficaciousness to a specific kind of 'physical causality', even if we had agreed on what it is. Thus, we are not in a position to claim that instantiations of apparent non-physical causation *inherit* 'nothing more than' the causal powers of the physical basis, because we do not know what such powers ultimately are. We may want to subscribe to the principle of the conservation of energy, but this does not entail any specific character of efficaciousness: nothing beyond 'energy' is inherited through 'causal inheritance'.

3.4 Is There Any Room for 'Epiphenomenalism' in a Monistic Ontology?

As we saw, according to Kim, the *conceptual* reduction, required by an accomplished reductionism, must be able not just to show *that* physical realisers perform a functional task, but to *explain how* they do it. And how is that supposed to happen? With reference to his favourite example of the reduction of genes to DNA, Kim states that «we have a theory that explains, *at least in broad but persuasive outlines*, how the DNA molecule is able to perform this causal task»,[31] or, in another occasion, that «[w]e *presumably have a story* at the microbiological level about how DNA molecules manage to code and

29 Kim 1999, 17-18.

30 Incidentally, Kim does not deny that macroproperties can be causally efficacious, or that there are effects of macroproperties, which are not to be straightforwardly found at the microlevel. But, according to him, macroproperties can be regarded as causally efficacious just because they are *themselves physical or reducible to physical*.

31 Kim 1997b, 51; my italics.

transmit genetic information».[32] The vagueness of these expressions is probably meant as a suggestion that the relevant scientific knowledge is still in progress. However, Kim's vague statements unwittingly reveal the weakness of his claim. Regardless of how closely we would look at the relevant microbiological description, we would never meet any spot where, so to say, the causal sequence 'displays its interiors', showing its transformation from physical-chemical regularity into the 'function to intergenerationally transmit phenotypic traits'. If at a certain level of detail we have regular connections of physical-chemical events, no increase in the level of detail will upgrade our description beyond physical-chemical correlations to the 'reasons why' the relevant *function* is performed. To expect that a greater detail in the same form of causal description could lead to a different level of *meaning*, where functional reasons are at home, is mere wishful thinking. An increase in the detail of the scientific knowledge of causal relations can be greatly important, since it provides knowledge of further elements on which we could in principle intervene to manipulate causal outcomes. But regardless of how detailed this analysis is, we are never in a position to discover anything like *reasons why* phenomena of a kind generate phenomena of another kind. The fact that life, as we know it, supervenes on a carbon-based material organisation does not, and never will, enable us to *deduce* the supervenient properties of the living from the subvenient properties of carbon and the elements it bonds to. The qualifying properties of the living are *phenomenal* and we have cognitive access to them by a route that is completely indifferent to any knowledge of the causal chains that may or may not realize them.

Now, according to Kim, if the way in which mental (intentional) properties produce their effects cannot be reduced to physical causality, we should accept that they are *epiphenomena.* But here we stumble into another mystery: what could ever be an *epiphenomenon* in a monistic ontology? In a monistic framework *even a delusional property* is a property that belongs to reality and in principle we must account for its existence. The possible delusional aspects could just concern the *actuality* of the *powers* that we attribute to some phenomena, but never the *existence of phenomena* as such.

This point refers us back to the starting point of phenomenological reflection. Anything that prima facie appears necessarily belongs to Being and can be never dismissed as irrelevant. Phenomena are Being. Mistakes can be done in attributing wrong *causal* connections to some phenomena, but not in the description of their apparent relations as such. Thus, we can wrongly believe that the experienced flash causes the experienced thunder; but the distinct

32 Kim 1999, 11; my italics.

experiences of the flash, the thunder plus the impression of a causal connection must all have a place in our ontology, and any improvement of our accounts of their actual connection will have to rely on those primal phenomena.

One could believe that granting that some phenomena may not have the causal powers that they appear to have is all we need to dismiss them as epiphenomena and to consider them irrelevant. But this is a hasty conclusion. We may not know, for instance, whether our will, desire, pain, etc. have all the effects that we attribute to them, but we do know that will, desire, pain, meaningfulness, hope, empathy, etc. do have an ontological status, that is, they have *efficaciousness* (at the very least, we should grant them the power to *mislead* us). Moreover, we also know that their *appearance* requires an explanation and precisely an explanation able to *account for their characteristic features* (the way they feel, the way they regularly connect with such and such experiences, etc.). But, if this is the case, we also know for certain that any ontology we may like to devise *must possess enough flexibility and power to make all those phenomena intelligible*. This is no negligible claim. This means that we cannot expect that our ultimate ontological categorisation will contain only contents homogenous with current physical language. And actually, we can extend this inadequacy to all natural sciences. This straightforwardly leads to the following conclusion: *no physicalistic (naturalistic) ontology can ever be true*, since it does not have the conceptual resources to explain the characteristic traits of will, desire, meaning, expectation, hope, intentionality in general, etc. We may discover regularities between events described in a physical, chemical, biological language and events described in mental terms. We may also discover a theoretical model (like evolutionism) that justifies the persistence over time of mental features. But we cannot make them intelligible by using only that semantic subset of natural language which is the language of physical laws, chemical formulas, etc. No ontology can be expressed in a language that is significantly less rich than a full-fledged natural language.

3.5 Second Intermediate Conclusion

In the previous section we have seen that there is no rational ground to support the idea that Being is essentially physical and that ontological efficaciousness must fit a specific model of physical efficient causation. In the present section, we have seen that we cannot reduce intentional relations (inclusive of functions) to extensional ones and that we cannot treat phenomena that are irreducible to naturalistic language as epiphenomena. All phenomena (appearances) have ontological status and any ontological account must take

seriously the phenomenological respect for phenomena as such. The main conclusion that we have to draw from this background is that there is no chance that a physicalistic (naturalistic) ontology be true: any view of Being that claims to explain away the qualitative plurality of appearances is committed to the subsistence of ontologically empty epiphenomena, which is just untenable. On the contrary, any tenable ontology must make room for each nuance of meaning, however slight. It may seem to me that an eclipse signals Divine wrath, a blooming cherry-tree vaticinates good luck or a smile of my beloved one foretastes eternal blessing. And I may be badly wrong about each of the guessed causal ties. But in any case it must be granted that precisely those meaningful contents that I entertained, as such, are part of Being, and that they must have an efficaciousness of their own (delusions may be less *successful* than sound scientific forecasts, but they possess no less efficaciousness).

Come to this point, any talk of the supervenience of the mental on the physical seems unacceptable, because there is no reason to treat the physical and the mental as pertaining to different levels of Being, respectively a base level and a peripheral (possibly epiphenomenal) level.

In conclusion, we have still to tackle the core phenomenon that suggests supervenience, that is, the *apparent* causal dependence of the mental on the physical.

4. Supervenience and Causation

4.1 What Do We Mean by 'Consciousness' ('Mind')?

When Davidson was speaking about the mental sphere he equated it with the *psychological* sphere. But under such a description it is by definition impossible to regard the mental as anything more than the attribute of a pretty limited set of material (biological) entities. This is the reason for making preferential reference to the phenomenological notion of *consciousness*, which rejects any a priori specific objectification. We should reject the idea that consciousness is just a biological faculty of some natural species, because this kind of interpretation discounts the nature of consciousness. In phenomenological terms consciousness is not a place, nor a thing, nor a container.[33] It is rather a *condition of manifestation* and an *ordering principle*. In phenomenological terms

33 Cf. BROUGH 2008.

consciousness cannot be reduced to either *reflective awareness* or *phenomenal consciousness*; and a fortiori consciousness cannot be reduced to alleged conscious behavior.

What is manifest ('conscious') to consciousness is always only a tiny emergent section of a thread whose articulations are as vast as our embodied life and all its settled experiences. We can bring to light this thread only through a careful descriptive and reflective process, which shows that there is remarkably more to consciousness than what is occasionally present in consciousness. By investigating the 'contents' of consciousness we meet continuous processes of co-ordination between our sentient body and its environment. We may not be aware of breathing, but we have tacit expectations (protentions) relative to it and we can modulate the rhythm of breathing according to tasks, emotions, etc. We may not be aware of how we co-ordinate distance and dimensions of perspectival objects, but we have learned to do it and, if needed, we can learn to correct it. One of the best-known contributions of phenomenological tradition is the recognition that consciousness is constitutively *intentional* and *embodied*; therefore, it cannot be fruitfully discussed without taking into account the activity and dispositions of our living body (*Leib*) and the co-ordination with its environment (*Umwelt*). This implies that whenever we talk about consciousness, we mention more than the object of ordinary 'psychology': we name *life* that issues into *intentionality,* and is recapitulated by *intentionality*.

The reference to consciousness is the reference to the *ubiquity* of intentional relations in 'the world'. Nothing can be said, referred to, evoked, perceived, felt, discerned without implying consciousness as a condition of manifestation and ordering principle.

The idea that we can treat the objective determinations of Being as subjectless self-standing items is a delusion, and being a rather fashionable delusion, does not make it less such.

4.2 Phenomenology of Efficient Causes

In Husserlian terms *reality* can be spelled as spatiotemporal existence (*real)* or as temporal existence (*reell*): physical entities exist in space and time and are therefore *real*, whereas all phenomena exist in time and are therefore *reell*. According to Husserl the sphere of spatiotemporal reality (*Realität*) and the sphere of efficient causality (*Kausalität*) belong together. Real (*real*) qualities are constitutively causal, in the sense that to *know a real X* means to have knowledge, consistent with experience, about how X behaves when it is

pushed, pressed, bent, heated, etc.[34] The same could be claimed about the reality of subatomic particles, magnetic fields or anything that we may regard as spatiotemporally real. This means that all attributes of 'naturalistic nature' emerge from the practices, tests, reactions that we deem worth considering. Up to this point, this position is not dissimilar from the 'operational realism' that we can find, for instance, in Ian Hacking's *Representing and Intervening*.[35]

But this picture must be completed on the *reell* side. In all interventions and interferences that bring to light *Realität* consciousness is pervasively present. Consciousness appears as 'complement' of whatever transcends consciousness. Consciousness is a steady negotiation with the transcendent sphere. And efficient causality turns out to be the prevalent conceptual tool to grasp such a negotiation.

Thus, we must make room for a complementary kind of efficaciousness which is not efficient causality, since there must be a kind of efficaciousness which is *able to bring to light efficient causes*. This is what Husserl calls '*motivation*'.[36] Yet, Husserl's account of the ontological status of motivation is less explicit than we may wish and we shall try to clarify this point by devoting a specific analysis to it.

Let us investigate what is involved in our understanding of efficient causality. At the very least, efficient causality is a form of *efficaciousness* which is supposed: α) to take place between *well-determined events*, which also define its 'direction'; and β) to follow a *temporal order* according to which antecedents cause successors.

α) *On well-determined events*. Efficient causality is meant to take place between circumscribable events in space and time, which produce effects on separate circumscribable events. When we meet complex or 'fuzzy' interactions we assume them to be the sum of linear interactions connecting well-determined events. Yet, there is something peculiar in this way to conceive efficaciousness. To begin with, there is nothing like a 'causal line' that neatly connects two events: each effect propagates in an indefinite plurality of ways, which are then regarded from time to time as 'side-effects', 'attritions', unintended, uninteresting or unknown implications, etc. Secondarily, no instance of efficient causation has identity in itself, because it intrinsically has neither *beginning* nor *end*. When we conceive of a specific event as a *cause*, we give a specific privilege to it. Such an event becomes the *active* side in a process where in point of fact

34 Hua IV, 45.
35 Hacking 1983.
36 Hua VI, 222; Hua IV, 132.

nothing deserves to be objectively considered either active or passive. Indeed, in the couple 'cause-effect', the former plays the role of the *agent*, while effect is posited as the passive side. Something can be a cause in this sense only if it is understood as an *irreducible starting point*, otherwise it would be just an occurrence in an infinite formless chain. But, nothing corresponds to an irreducible origin in natural processes *per se*. The mere regularity of successions would never entitle us to say that the antecedent event (A) *causes* the constantly associated follower (B). This becomes clear if we notice that any constant succession of two elements could be always alternatively explained by a third causal factor (C) preceding both A and B and producing them separately, without them directly connecting at all.[37] From this remark, Von Wright drew the correct conclusion that causality is unintelligible without reference to the notion of *action*, which is required in order to conceive instances of efficaciousness starting from A and directed to B, beyond mere association. Causal sequences do not begin, do not end and do not go in any single direction; only if and when events are actively circumscribed by us, origins, ends and directions may emerge. Moreover, there is no originating event in a causal sequence also for a second essential reason: 'objectively' there is no such a thing as a circumscribed *event* at all. In themselves *events have no specifiable boundaries*, because there is *no intrinsic boundary between what happens and what is caused to happen* (in all conceivable forms). There is no way to determine *iuxta propria principia* that an event finishes and another one begins. We can and do determine events, but this is something that can happen only in the wake of our *motivations*.

β) *On temporal ordering.* The Husserlian reflection on time and consciousness can provide a radical perspective from which to interpret the meaning of temporal ordering in efficient causality. Past and future are no events in the world, but, according to Husserl (and previously Kant), they are preconditions for events to be intelligible. The Husserlian analysis primarily tackles the most fundamental dimension of temporality, that is, temporality as it is immediately experienced in *perception*. In its wake emerges the constitution of secondary temporality (the past and future of which we are reflectively aware).[38] As we know, the perceptual dimension can be never 'bypassed', because the phenomena of the perceptual sphere are the ones that allow all theories (representations) to be formulated and all empirical hypotheses to be tested.

Perception gives shape to meaningful entities only insofar as a current

37 Von Wright 1971, 73-4.

38 Hua X, 45.

impression can be read in the wake of the *retention* of the just flown sequence of impressions.[39] 'Retention' names the train of impressions that is retained in consciousness whenever we are engaged in intentional acts (primarily perceptual ones, but not only). Retentions are given to *thematic* consciousness (attention) only secondarily, by a reflective act, while primarily they have a 'latent' unfocussed nature. We could never hear a piece of music as a piece of music if our perception referred just to the instantaneous presence of the sound: retentions allow us to understand a musical note as 'coming from the preceding ones' and this is what confers musical meaning. Yet, retentions are not 'containers' of impressions, since they also retain the sensuously *empty* intervals as such, with their duration and order. This happens not just for acoustic phenomena, but for all perceptual acts, because all perceptual apprehension is essentially diachronic (the currently visible surface of the perceived house 'adumbrates' its invisible sides as available for further exploration).

But retentions are not enough to give shape to primal temporality. The picture must be completed with the horizon of tacit expectations, which Husserl calls '*protentions*'. Protentions can be brought to evidence by noting that we may be *surprised* by some events, even if we did not consciously expect a different event. In these cases we realize that an immediate tacit expectation was latently at work. I may not have any thematic expectation about the solidity of the ground, or the smooth working of my respiration, but I would be nevertheless surprised if the ground would disappear under my feet, or my next breath failed; and this surprise shows that I had qualified tacit expectations (protentions) about those situations. (I had them next to countless further unnoticed protentions).

A superficial reading of Husserl's pages on inner time-consciousness can lead to psychologistic interpretations where retention and protention may be interpreted as psychological faculties (for instance, retention would be just another name for short-term memory). This approach would trivialize Husserl's analysis, by depriving it of its ontological grip. When Husserl speaks about the *constitution*[40] *of temporality*, he is not speaking of mere psychological

39 HUA X, 38.

40 The phenomenological term constitution (*Konstitution*) does not mean 'creation', but a process of negotiation between consciousness and transcendence, which issues in the bestowal of sense (meaning and order) to the world. The process of constitution is not akin to the sovereign activity of subjective idealism because at the roots of constitution there is always a *passive* genesis (HUA XI, 117-145) and because in the process of constitution the subject itself does not remain untouched (HUA XV, 283-284, 546). Thus, consciousness relates to its objects as an *essential constituting correlation* (HUA IV, 179; HUA VI, 169).

temporality, while temporality 'in itself' would lie somewhere else, untouched by consciousness. Transcendent events (changes) may well happen independently of consciousness, but the *order* of succession of events is conceivable only insofar as there is a constituting consciousness (which is *not* identical with a physical brain).

Let us formulate the supposition that there is an intrinsic *order of succession* in nature, which can be *recorded* by consciousness, but does not need any consciousness to subsist. Now, is there any physical description that can account for the establishment of such an order of succession? Let us think of an ancient celestial body, like the moon. We believe that its craters are due to the impact over time of many meteorites. But, in the absence of any interpretive activity, in each 'moment' of cosmic unfoldment the bygone succession of collisions exists only as a finite present amount of signs or traces left by the collisions. And actually we could not even mention *signs* and *traces,* in the absence of an interpreting consciousness. At most we might have something like a snapshot of the physical state of the meteorite. The surface of the moon would be like an impressed film where all impressions have taken place on the same photogram: nothing can tell us what the order of succession of the impressions was, nor if any succession took place at all. But let us go one step further. Let us unburden consciousness from the need to produce a retrospective theory about craters and meteorites and let us imagine that we have a faithful footage of the history of the collisions of meteorites. Now, in the absence of any instantiation of consciousness this would be just another present piece of matter (an exposed photographic film, say), which does not bear the slightest intrinsic reference to any order of succession of events. We had just a further piece of matter in present existence. Indeed, to become indicative of an order of succession the film should be *unfolded and perceived*, and this process requires temporal consciousness.

But, let us go a final step further. Let us imagine that we ourselves had personally perceived the natural history of the moon, thus, instead of the surface of the meteorite or the filmed record, we had *a present cerebral state*: are there reasons to believe that things would be any different? If our film was no longer a film, but a physical trace recorded on our cerebral hardware, would it make any difference? The point is that no physical description of a present amount of matter, whatever its configuration, can be conceived as repository of an order of succession. Present events can become *signs* or *traces* of something only if they are taken up by a *living consciousness*. This straightforwardly implies that there is no temporality, as a specific order of succession, in the absence of consciousness. The living present retains the just gone events *in*

sight of the coming ones, and expects specific events to come in *the wake of* the retained ones: this living process weaves together experiences and constitutes meaningful units unfolding over time.

Now, we can come back to efficient causality. From what we have just seen we must say that efficient causality is neither 'invented' by consciousness, nor passively 'found'. The efficaciousness of nature, its spontaneous power, its independent changes are transcendent instances and are also features of efficient causality. But the articulation in parts (events) of causality, its direction and its order of succession require a non arbitrary 'selection' and 'synthesis' by consciousness, without which the features that pre-eminently characterize efficient causality would not subsist.

This is an essential point because the traits of efficient causality that we consider incompatible with subjectivity are precisely the ones that appear as *complementary to* specific *motivations*. Efficient causality is one of the main forms taken by *agency's endeavor to gain power on transcendent reality and govern it*. The cause is conceived always already *in sight* of its effect. If we tried to conceive of naturalistic causes as they are supposed to objectively be, that is, as pure subjectless self-standing reality, they would be devoid of any identity and direction (in space and time), and therefore they could never support predictions and rational laws (among which the laws of Davidson's second principle).

To sum up: the motivational dimension is absent from efficient causality precisely because the aim that *motivates* the identification of efficient causes is the discovery of *means* for an (ideal) agent's ends. Efficient causality does not contemplate finality and meaningfulness because finality and meaningfulness are always already at work in the constitution and use of efficient causes. It is the most curious misunderstanding the one that makes us look in vain for instances of mental causation in the midst of physical relations, without seeing that the very essence of causality in nature is determined by its role for motivated agents.

4.3 On the Efficaciousness of Consciousness

The picture that has begun to emerge is such that the apparent ubiquity of efficient causality in nature must be complemented by the ubiquity of consciousness (not as 'state of mind', but as structural[41] condition). Yet, this

41 The appropriate term here would be '*transcendental*' rather than 'structural'. Consciousness is a transcendental condition of the actual and potential manifestation of all ontological determinations. But the term 'transcendental' is encumbered by a complex and occasionally controversial tradition, a reference to which may run the risk of being more misleading than illuminating. Therefore we

thesis can be easily misinterpreted and now it is time to try a conclusive clarification.

To begin with, let us repeat that we are not saying that consciousness is an independent source of ontological efficaciousness: this would be tantamount to dualism. We rather claim that monistic ontological efficaciousness has a plurality of forms of manifestation, which are poorly captured by the exclusive reference to efficient causality. Efficient causality is just the most effective form to organize efficaciousness in order to make it functional for agency.

Secondly, the ubiquity of consciousness in the determinations of nature does not imply anything like a 'priority of final causation'. 'Final causation', if by that we mean the idea that future events are causally efficacious on present ones, is a spurious notion. We cannot conceive of ends exercising any direct causal power from the future. Logical paradoxes can be easily produced, if we assume that the future *causally* affects the present.[42] We are *motivated* by absent options, by desired situations, by open possibilities, by the 'will to survive', etc. but whenever such motivations are at work, their efficaciousness is anchored in a present state of affairs, exactly as efficient causality is. Just, efficient causality is not what it seems, since it is always intrinsically inhabited by an essential motivational dimension.

But, if this is the case, how is the peculiar efficaciousness of consciousness to be conceived? The answer can be set forth in two stages.

4.3.1 Consciousness as Selective Form

The first and most elementary level at which the efficaciousness of consciousness can be acknowledged is as *selective form* (or *essence*). We are entitled to think of such a selective 'power' in terms of embodied preferences (needs, desires, etc.), but we should recall here the extended sense of consciousness that was previously mentioned. On a monistic background we must take for granted that *consciousness belongs to Being*. If we keep in mind that we do not know what *matter* ultimately is (*i.e.*, what its *properties* ultimately are), we could substitute 'matter' for Being and render this point through a simplified expression like: 'consciousness inheres in matter'.

But, if this is the case, we cannot think of consciousness as something that 'acts' *on* matter from without, because consciousness does not exist outside matter. Activity implies a distance between the agent and what is acted on, which is here impossible. This is why we suggest to read consciousness as *selective form*: to be active is originally a *way of being* and the 'selection' that

have preferred not to rely on such a term throughout the paper.

42 Cf. BLACK 1956.

consciousness *does* coincides with what consciousness *is*.

How could we conceive of consciousness which 'selects' within Being? First, we must conceive of selection not as an activity, but as something like an essential property. Secondly, such selection is no unilateral power on what is selected, but must be conceived as a pole of *correlation*. Thirdly, if we introduce the temporal dimension, we should conceive of such a selection as something that *manages to persist*: an *enduring correlation*.

All that may sound abstract and cryptic, but it may become clearer if we envision it in ordinary naturalistic terms. Let us summon a common account of natural history, and of natural selection in it. Here, it would not sound strange to describe consciousness as something that emerges in biological life, and life as something that emerges from matter and develops by being selected for survival. From this perspective, nothing is more straightforward than saying that consciousness (life) potentially inheres in matter, and that consciousness (life) exists thanks to selection.

Yet, some caution is needed. We are probably inclined to think of selection as something that *happens to life*. But in a monistic framework we are not entitled to separate what happens to consciousness (life) from what happens to Being. What we should say is that selection happens to Being and therefore to life. Selection is an ontological process that brings to light consciousness (life) and simultaneously brings to light the complementary reality (*Umwelt*) where consciousness (life) can exist as an enduring pole of correlation. In the absence of the selecting/selected principle that we call consciousness (life) we could not take hold of any point of view in the universe and thus we could not establish any meaningful unit in the universe. We could never say that a thing has such and such parts, a motion such and such direction, an event such and such extension in space and time, etc. without occupying a point of view from which to carve out our descriptions. In the absence of any selective point of view there literally are no things, no directions, no events (at most, we can say that 'in itself' there is just their ideal detectability).

Thus, consciousness (life) is not to be read as the mere outcome of selection, but also as the protagonist of selection: it gives shape to (its) reality by giving shape to itself (and this is something that is done neither *by* consciousness, nor *to* consciousness).

4.3.2 Supervenience Versus Formal Causality

After this change of perspective, we can tackle the issue of the alleged 'supervenience of the mental on the physical'. In normal action and perception

(as in Davidson's first principle) we seem to deal with two alternative causal directions, which may be taken to support two different forms of causality (efficient, or event causality and mental, or agent causality). Yet, there are cases where this distinction seems to collapse.

If I am drowsy and sluggish, I may drink a coffee, a tea or even take some more specific chemical substance, and this may clear my mind, boost my will and increase my vigilance. If, as Husserl writes,[43] I ingest santonin, my chromatic perception can be affected and I may happen to see everything yellowish (xanthopsia). If I undergo cortical electrostimulation of the occipital lobe, I may happen to perceive phosphenes or other visual hallucinations. Examples could be indefinitely multiplied. In all these cases nothing like an intentional act is in place. Both, perception and action are intentional acts (in phenomenological sense), but here we are apparently in front of a non-intentional dependence of mental events from physical events: this kind of dependence motivated the appeal to supervenience. But after our refusal of the physicalistic ontology and its replacement with a phenomenologically based quasi-Spinozian vision, how can we capture such an apparent dependence of mental states on physical causes?

Let us go back to Kim's exclusion argument and his criticism of 'mental causation'. Kim's argument relied on the idea that there was a specific way of producing effects by physical events (efficient causality) and that complex material organizations (like the brain) could just inherit their causes from lower physical levels (*e.g.*, micro-physics). As we saw, there is no account of 'physical causation' that could exclude reference to intentionality and be preserved at different levels. The only 'identity of physical causation' which may be apt to be preserved is described by the principle of the conservation of energy.

In order to see how consciousness may have efficaciousness without violating the conservation of energy, we must first notice that all experience (scientific or not, under *Epoché* or not) shows the ubiquitous subsistence of *selective thresholds* and *discontinuities* in efficacious relations. That is, not everything that happens or changes at a certain level, or in a certain entity, produces proportional events or changes at another level, or generally outside the relevant entity. We recognise, for instance, that not everything that *happens in* an atom, or a cell, or a planet, etc. produces effects outside the atom, the cell, and the planet respectively. Indeed, it is precisely such kind of discontinuity that allows speaking of the threshold between 'inside' and 'outside' of the relevant items. We make experience of organisms which change physical constituents over time without losing identity; we detect transitions of energy levels in atoms

43 HUA IV, 63f.

that remain the same atoms; we identify molecular oscillations, without molecules turning into different molecules. In a world devoid of such discontinuities each (putative) elementary event should fluidly and continuously spread its effects without boundaries and no entity endowed with identity could *appear*. Such thresholds could be considered the intuitive correlate of what physics has fruitfully explored under the category of *quanta*; yet, any physical technicality is here beyond the point. The core question is only that the conservation of energy is perfectly compatible with *essential discontinuities* in the transmission of effects. This, in turn, implies possible asymmetries of causal transmission, whenever a propagating effect comes to a *qualified* threshold: small changes that reach specific thresholds can spark macroscopic effects, big changes unable to prime a specific threshold can see their effects dissipated.[44] This means that, when we question which specific changes are brought about by other specific changes, the reference to the conservation of energy is quite unhelpful as to the possibility to define the *nature* of determinate effects. Whatever the quantity of energy available in a precedent stage, it is essential to know what was the *nature* or *essence* of the elements involved. Some changes within some configurations can produce effects that the same changes outside of those configurations cannot, because there are thresholds that *select* and *modulate* already existing effects.[45] This point could be expressed by recovering from the philosophical tradition the notion of *formal causation*. Formal causation should not be primarily read as transferral of physical quantities, nor as succession of efficient causes and passive effects (although both descriptions are available): here the kind of efficaciousness simply depends on the *nature of the elements involved*. If we resort to the Aristotelian opposition between formal and material causes, we can say that the same ontological efficaciousness (or the same amount of energy) is like a constant material cause that produces different realities according to the 'form' that it takes.

How can we then interpret the above mentioned paradigmatic instances of supervenience? To begin with, when we grant that material interventions on cerebral matter produce effects in the mental sphere, we should refrain from thinking of them as causal determination of mental contents. The idea of a bottom-up causal continuity from physical cause to mental event denies that

44 These phenomenal traits have been focus of interest in recent discussions on emergent properties. Processes of *amplification*, like the ones commonly exploited by cloud chambers or Geiger counters, are often mentioned as a key to conceptualise emergence (Deacon 2006; Bitbol 2007). More generally, *non-linear processes* (Harth 2008, 61), with special reference to *positive or negative feedback* (Bickhard/Campbell 2000, 342) have been considered natural carriers of emergence.

45 Cf. Murphy 2006, 227.

the caused mental event may have a contribution of its own to further causal developments (no 'downward mental causation'). But we have to acknowledge that it is the nature of any specific concretion of Being (things, structures, etc.) to *select* the way in which transcendent causes can *affect the concretion itself* and simultaneously to *select which further effects can be transmitted*.

It may well be that technical improvements in neurophysiologic interventions will increasingly enable us to 'cheat' a mind through cerebral changes, and by that means to induce more and more sophisticated *illusions elaborated by the mind*. But, regardless of how sophisticated such interventions may be, they could never be interpreted as causal reductions of the mind. Any such 'cheating' presupposes the *living spontaneity* of the 'cheated' mind, which, insofar as it is a mind, *is motivated* to *deal with* the raw matter of 'bottom-up effects', while being irreducible to them.

The essential point here is that not all physical causes operating on the body (or the brain) can affect the mind *as mind*. Some physical causes can kill life and *annihilate* the mind: they do not issue into mental events. Some other physical causes (from solar winds to fleeting quarks) do not interfere with the mind and remain *unperceived*: they do not issue into mental events either. Some physical causes can produce 'interferences' that cannot be *synthesised* by the mind, while being noticed: confusion, tiredness, excitement, etc. may be causally produced without obtaining the status of intentional object. And finally, some physical causes do become intentional objects positioned in the world: this is pre-eminently the case of 'inner' and 'outer' percepts (inclusive of illusions).

5. Conclusion

What we call 'activity of the mind' should be primarily read as the selective and modulating *nature* of the mind. We must not introduce discussions about 'mental causation' starting from what the mind *does* or can do, but from what the mind *is*. If we accept that there is no unitary essence of causation and if we make room for the intuition that ontological efficaciousness is qualified by the selective character of *ontological thresholds*, then we can see why the idea of the supervenience of the mind on matter is misleading. The mind *qualifies* causation in ways that can be said to be *potentially inherent in matter*.

Consciousness does not produce any efficaciousness (energy) of its own and in this sense no violation of the principles of conservation is envisaged. Consciousness brings to light an ordered world by having emerged, and existing, as ontological sphere co-ordinated with (part of) the world (*Umwelt*). Thus, we can say that consciousness depends on 'matter', but not that it

depends on physical properties (*i.e.*, physical ways of producing effects). This means that when we intervene on matter according to our physical categorization of it, we may affect consciousness, but we do not turn the essence of consciousness (preference, feeling, spontaneity, etc.) into something else (mechanism, chaos, meaninglessness, etc.). We can switch off consciousness, we can boost or obstacle the working of consciousness, but we must rely on its weak and subtle abilities in order to produce 'mental events'. At the same time, such weak and subtle abilities, when they operate in environments with which they are well co-ordinated, can have vast and enduring effects.

References

BICKHARD, M. AND CAMPBELL D. T., 2000. «Emergence», in ANDERSON, P., EMMECHE, C., FINNEMANN, N., AND CHRISTIANSEN, P. (eds), *Downward Causation: Minds, Bodies, and Matter,* Aarhus: Aarhus University Press, pp. 322-348.

BITBOL, M. 2007. «Ontology, matter and emergence». *Phenomenology and the Cognitive Sciences*, 6, pp. 293–307.

BLACK, M. 1956. «Why Cannot an Effect Precede Its Cause». *Analysis*, 16/16, pp. 49-58.

BLOCK, N. 2007. *Consciousness, functionalism, and representation*. Cambridge MA/ London: The MIT Press.

BROUGH J.B. 2008. «Consciousness is not a Bag: Immanence, Transcendence, and Constitution in *The Idea of Phenomenology*». *Husserl Studies*, 24, pp. 177-191.

DAVIDSON, D. 1970. «Mental Events», in *Essays on Actions and Events*, Oxford: Clarendon Press, pp. 207-227.

DEACON, T. 2006. «Emergence: The Hole at the Wheel's Hub», in Clayton, P. and Davies, P.S. (eds.), *The Re-Emergence of Emergence.* Oxford: Oxford University Press, pp. 111–150.

DOWE, P. 2000. *Physical Causation*. Cambridge: Cambridge University Press

Fair, D. 1979. «Causation and the Flow of Energy». *Erkenntnis*, 63, pp. 219-250.

HACKING, I. 1983. *Representing and Intervening.* Cambridge: Cambridge University Press.

HARTH, E. 2008. «The Element of Time in the Emergence of Mental Phenomena». *Journal of Consciousness Studies*, 15/4, pp. 54–65.

HUSSERL, E. 1952. *Ideen zu einer reinen Phänomenologie und phänomenologischen*

Philosophie. *Zweites Buch*, Husserliana Bd. IV, Den Haag: Martinus Nijhoff. (Quoted as: HUA IV)

HUSSERL, E. 1954. *Die Krisis der Europäischen Wissenschaften und die transzendentale Phänomenologie*, Husserliana Bd. VI, Den Haag: Martinus Nijhoff. (Quoted as: HUA VI)

HUSSERL, E. 1966. *Vorlesungen zur Phänomenologie des inneren Zeitbewusstseins (1893-1917)*, Husserliana Bd. X, Den Haag: Martinus Nijhoff. (Quoted as: HUA X)

HUSSERL, E. 1966. *Analysen zur passiven Synthesis. Aus Vorlesungs- und Forschungsmanuskripten 1918-1926*, Husserliana Bd. XI, Den Haag: Martinus Nijhoff. (Quoted as: HUA XI)

HUSSERL, E. 1973. *Zur Phänomenologie der Intersubjektivität. Dritter Teil: 1929-1935*, Husserliana Bd. XV, Den Haag: Martinus Nijhoff. (Quoted as: HUA XV)

HUSSERL, E. 1973. *Die Idee der Phänomenologie. Fünf Vorlesungen,* Husserliana Bd. II, Den Haag: Martinus Nijhoff. (Quoted as: HUA II)

Husserl, E. 1976. *Ideen zu einer reinen Phänomenologie und einer phänomenologischen Philosophie.* Husserliana Bd. III/1, Den Haag: Martinus Nijhoff (Quoted as: HUA III/1).

HUSSERL, E. 2006. *Späte Texte über Zeitkonstitution (1929-1934). Die C-Manuskripte*, Husserliana Materialien Bd. VIII, Dordrecht: Springer. (Quoted as: HUAMAT VIII).

KIM, J. 1992. «'Downward Causation' in Emergentism and Nonreductive Physicalism», in Beckermann, A., Flohr, H. and Kim, J. (eds.) *Emergence or Reduction?* Berlin: De Gruyter, pp. 119-138.

KIM, J. 1997a. «Does the Problem of Mental Causation Generalize?». *Proceedings of the Aristotelian Society*, 97, pp. 281-297.

KIM, J. 1997b. «Explanation, Prediction, and Reduction in Emergentism». *Intellectica,* 2/25, pp. 45-57.

KIM, J. 1999. «Making Sense of Emergence». *Philosophical Studies*, 95, pp. 3-36.

KIM, J. 2005. *Physicalism, or Something Near Enough.* Princeton: Princeton University Press,.

KIM, J. 2006. «Being Realistic about Emergence», in Clayton, P. and Davies, P.S. (eds.), *The Re-Emergence of Emergence.* Oxford: Oxford University Press, pp. 189-202.

MURPHY, N. 2006. «Emergence and Mental Causation», in Clayton, P. and Davies, P.S. (eds.), *The Re-Emergence of Emergence.* Oxford: Oxford University Press, pp. 227-43.

NAGEL, E. 1961. *The Structure of Science*. New York: Harcourt, Brace and World.

PAPINEAU, D. 2002. *Thinking about Consciousness*. Oxford: Clarendon Press.

SALMON, W. 1998. *Causality and Explanation.* New York: Oxford University Press.

VON WRIGHT, G.H. 1971. *Explanation and Understanding*. London: Routledge.